

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-275279

(43)Date of publication of application : 18.10.1996

(51)Int.Cl. H04R 1/00
G10L 3/00
G10L 3/00
H04R 3/00

(21)Application number : 07-075875 (71)Applicant : FOSTER ELECTRIC CO
LTD

(22)Date of filing : 31.03.1995 (72)Inventor : YAMAGUCHI TAKURO

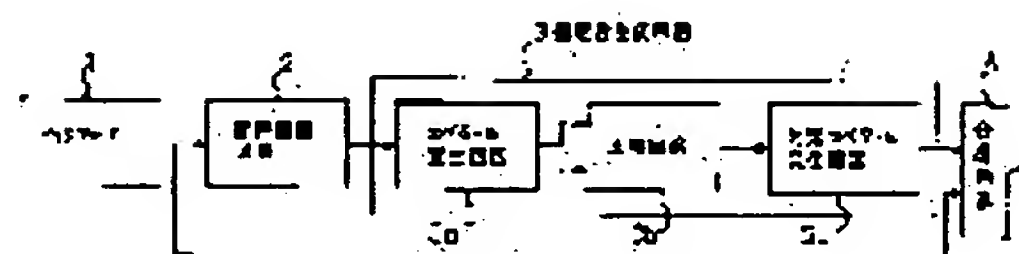
(54) VOICE PICKUP SYSTEM

(57)Abstract:

PURPOSE: To realize the voice pickup system in which a voice is sent with a high articulation without picking up a surrounding noise.

CONSTITUTION: The voice pickup system detecting a bone conduction sound or an air conduction sound by a pickup 1 is provided with a voice recognition means 2 capable of recognizing a sound detected by the pickup 1 and a spectrum generating means 3a using the sound data recognized by the voice recognition means 2 to generate a spectrum with a frequency component when the sound data recognized with a conventional microcophone, and also with a comparison means 3b

comparing the sound spectrum detected by the pickup 1 with the generated spectrum to obtain a missing frequency component, a missing spectrum generating means 3c generating a sound of the frequency component missing in the sound detected by the pickup 1 as a supplement sound according to the result of comparison by the comparator means 3b, and a synthesis means 4 synthesizing the sound detected by the pickup 1 with the supplement sound generated by the missing spectrum generating means 3c and providing the synthesized output.



LEGAL STATUS

[Date of request for examination] 12.02.2002

[Date of sending the examiner's decision]

of rejection]

[Kind of final disposal of application
other than the examiner's decision of
rejection or application converted
registration]

[Date of final disposal for application]

[Patent number] 3647499

[Date of registration] 18.02.2005

[Number of appeal against examiner's
decision of rejection]

[Date of requesting appeal against
examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

JPO and NCIP are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.**** shows the word which can not be translated.

3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] In the voice pickup system which detects a bone conduction sound or a respiratory tract sound by pickup The data of the sound recognized with a speech recognition means by which the sound detected by said pickup can be recognized, and said speech recognition means are used. A spectrum generating means to generate the spectrum of a frequency component when the usual microphone detects the data of this recognized sound, A comparison means to ask for the frequency component which compares the spectrum which this spectrum generating means generated with the spectrum of the sound detected by said pickup, and is missing, A lack spectrum generating means to generate the sound of the frequency component which the sound detected by said pickup from the result of a comparison with said comparison means lacks as a complement sound, The voice pickup system characterized by having a synthetic means to compound and output the sound detected by said pickup, and the complement sound generated with said lack spectrum generating means.

[Claim 2] The voice pickup system characterized by having a speech recognition means by which the sound which detected the bone conduction sound or the respiratory tract sound by said pickup in the voice pickup system detected by pickup can be recognized, and an artificial sound generation means to generate the sound corresponding to the data of the sound recognized with said speech recognition means as an artificial sound.

[Translation done.]

* NOTICES *

JPO and NCIP are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. **** shows the word which can not be translated.

3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to the voice pickup system which detects a bone conduction sound and a respiratory tract sound by pickup in more detail about a voice pickup system.

[0002]

[Description of the Prior Art] As a microphone used in the environment where the noise exists, the bone conduction sound pickup which detects a bone conduction sound is known.

[0003] This bone conduction sound pickup is stuck on a user's face and head, and vibration of the face and the head accompanying utterance of a user is detected. Moreover, pickup of a configuration like an earphone is inserted in an external ear as a similar microphone, and what detects the respiratory tract sound of external auditory meatus exists.

[0004] The pickup which detects the above bone conduction sounds and a respiratory tract sound has the advantage of being easy to detect a user's voice made into the purpose, without gathering the surrounding noise comparatively.

[0005]

[Problem(s) to be Solved by the Invention] However, when the spectrum of the usual voice is 300Hz – 3kHz, as shown in the drawing 5 property B, it has the problem to which a disregard level falls in a frequency domain 1kHz or more.

[0006] That is, as compared with the sound signal (drawing 5 property A) detected with the usual microphone, an audio low-pass component is emphasized, a high-frequency component becomes sensibility that it falls gradually (drawing 5 property B), and detection sounds of pickup of the format which detects voice indirectly, such as a bone conduction sound and a respiratory tract sound, have the problem to which articulation falls.

[0007] Drawing 6 is the property Fig. showing the frequency characteristics at the time of actually detecting male voice by bone conduction sound pickup. The situation of a fall of a high-frequency component can be read also in this property Fig. Although the degrees of a fall of a high-frequency component differ a little in fact the pressure at the time of suppressing bone conduction sound pickup on the face or a head, according to a male/woman, etc., there is no change in a high-

frequency component falling.

[0008] This invention was made in view of the above-mentioned trouble, and the purpose is in offering the voice pickup system which can transmit voice in the condition that articulation is high, without gathering the surrounding noise.

[0009]

[Means for Solving the Problem] The artificer of this application completes a header and this invention for the configuration which can transmit voice with high articulation to articulation also in the pickup which detects the bone conduction sound which had the problem, and a respiratory tract sound in respect of frequency characteristics conventionally, as a result of inquiring wholeheartedly that faults, such as articulation expected in the conventional voice pickup system, should be improved.

[0010] Therefore, this invention which is a means to solve a technical problem is constituted so that it may explain below.

(1) Namely, 1st means to solve the above-mentioned technical problem In the voice pickup system which detects a bone conduction sound or a respiratory tract sound by pickup The data of the sound recognized with a speech recognition means by which the sound detected by said pickup can be recognized, and said speech recognition means are used. A spectrum generating means to generate the spectrum of a frequency component when the usual microphone detects the data of this recognized sound, A comparison means to ask for the frequency component which compares the spectrum which this spectrum generating means generated with the spectrum of the sound detected by said pickup, and is missing, A lack spectrum generating means to generate the sound of the frequency component which the sound detected by said pickup from the result of a comparison with said comparison means lacks as a complement sound, It is the voice pickup system characterized by having a synthetic means to compound and output the sound detected by said pickup, and the complement sound generated with said lack spectrum generating means.

[0011] In addition, in such a voice pickup system, it is also possible to carry out study for speech synthesis to a system beforehand, and to consider as a lack spectrum required since the voice modeled on a user's voice is generated as a complement sound generated in a lack spectrum generating means.

[0012] (2) Moreover, 2nd means solve the above-mentioned technical problem is a voice pickup system characterized by to have a speech-recognition means to by which the sound which detected the bone-conduction sound or the respiratory tract sound by said pickup in the voice pickup system detected by pickup can be recognized, and an artificial sound generation means generate the sound corresponding to the data of the sound recognized with said speech-recognition means as an artificial sound.

[0013] In addition, in such a voice pickup system, it is also possible to also generate the voice which was made to carry out study for speech synthesis to a system beforehand, and was modeled on a user's voice as an artificial sound generated in an artificial sound generation means, and to generate another general voice.

[0014]

[Function] In the voice pickup system which is 1st means to solve a technical problem The sound which detected the bone conduction sound or the respiratory tract sound by pickup, and was detected by pickup is recognized with a speech recognition means. The spectrum of a frequency component when the usual microphone detects the data of this recognized sound is generated using the data of the sound recognized with the speech recognition means. Thus, it asks for the frequency component which compares the generated spectrum with the spectrum of the sound actually detected by pickup with a comparison means, and is missing. The sound which generated the sound of the frequency component which the sound detected by pickup from the result of a comparison with a comparison means lacks as a complement sound with the lack spectrum generating means, and was detected by pickup, and the complement sound generated with the lack spectrum generating means are compounded and outputted in a synthetic means. [0015] According to the above voice pickup systems, the clear sound signal near the usual voice can be generated by carrying out speech recognition and complementing a lack spectrum, after detecting a bone conduction sound and a respiratory tract sound, without spoiling the description of his voice. Moreover, it is not influenced of the surrounding noise.

[0016] In the voice pickup system which is 2nd means to solve a technical problem, an artificial sound generation means generates the sound corresponding to the sound which has recognized the sound which detected the bone conduction sound or the respiratory tract sound by pickup, and was detected by pickup with the speech recognition means, and has been recognized with the speech recognition means as an artificial sound.

[0017] According to the above voice pickup systems, the clear sound signal near the usual voice can be generated now by generating the artificial voice which detects a bone conduction sound and a respiratory tract sound, and corresponds with the data by which speech recognition was carried out. Moreover, it is not influenced of the surrounding noise.

[0018]

[Example] One example of this invention is explained to a detail using a drawing.

<The voice pickup structure of a system (1)> The voice pickup structure of a system which is one example of this invention is first explained using drawing 1 .

[0019] Pickup 1 is a detection means to detect a bone conduction sound or a respiratory tract sound, and a respiratory tract sound microphone corresponds as what detects a bone conduction microphone (bone conduction sound pickup) and a respiratory tract sound as what detects a bone conduction sound.

[0020] The speech recognition circuit 2 is a circuit which recognizes well-known voice or well-known syllable, and whether it learns each description part of a user or it does not perform such study, it is not cared about.

[0021] The complement sound generation circuit 3 is a kind of artificial sound generating circuit which generates the sound of the frequency component which the sound detected by pickup 1 lacks using the data of the sound (voice, syllable) recognized in said speech recognition circuit 2 as a complement sound. Moreover, this complement sound generation circuit 3 consists of spectrum generating circuit 3a, comparator circuit 3b, and lack spectrum generating circuit 3c.

[0022] Spectrum generating circuit 3a generates a spectrum when the recognized sound is detected by the usual microphone using the data of the sound recognized in said speech recognition circuit 2. Comparator circuit 3b compares the spectrum which said spectrum generating circuit 3a generated with the spectrum of the sound detected by said pickup. Lack spectrum generating circuit 3c generates the spectrum (lack spectrum) of a part according to the difference of a spectrum as a complement sound according to the comparison result of comparator circuit 3b.

[0023] The synthetic circuit 4 is an output means to compound and output the sound (a bone conduction sound, respiratory tract sound) detected by pickup 1, and the complement sound generated in the complement sound generation circuit 3.

Actuation of the voice pickup system which is one example of <actuation [of a voice pickup system] (1)> this invention is constituted by each step of **, **, **, **, and ** as roughly divided and shown below. Order is explained for this step later on.

[0024] ** Detection of a sound (a bone conduction sound, respiratory tract sound) : detect a bone conduction sound or a respiratory tract sound using pickup 1.

** Recognition of a sound (a bone conduction sound, respiratory tract sound) : recognize the detection sound of pickup 1 in the speech recognition circuit 2. In this case, you may be any although recognition by the single sound or recognition by syllable changes with recognition approaches of the speech recognition circuit 2. Moreover, it does not matter whether it learns and recognizes a user's voice, or it recognizes without performing study which specified the user.

[0025] ** Spectrum generating of a recognition sound : generate the spectrum (this is hereafter called spectrum of a standard sound) which will be obtained when spectrum generating circuit 3a detects with the usual microphone using the data of a recognition sound. For this reason, the spectrum which spectrum generating circuit 3a shall have the spectrum according to a recognition sound (a single sound, syllable), and corresponds according to a recognition sound is called.

[0026] As a spectrum in this case, you may have a spectrum for every user and may have a standard spectrum. Moreover, when it has a standard spectrum, you may make it have some spectrums like an adult male, an adult woman, and a child.

[0027] ** The spectrum comparison with a standard sound and a detection sound : compare the spectrum of the spectrum of a standard sound, and a detection sound in comparator circuit 3b.

[0028] For example, if it explains using drawing 5 , the lack spectrum C of a detection sound (= A-B) will be computed for every recognized detection sound by comparing the spectrum A of a standard sound with the spectrum B of a detection sound.

[0029] ** Complement sound (lack spectrum) generating : according to the data of the lack spectrum obtained by comparator circuit 3b, lack spectrum generating circuit 3c generates the signal of a lack spectrum. Also in this case, as a

generated lack spectrum, you may have a spectrum for every user and may have a standard spectrum like spectrum generating circuit 3a. Moreover, when it has a standard spectrum, you may make it have some spectrums like an adult male, an adult woman, and a child.

[0030] ** Composition with a detection sound and a complement sound : compound a lack spectrum and the spectrum of a detection sound in the synthetic circuit 4. By this synthetic processing, the lack spectrum of a detection sound is added as a complement sound, and the composite tone of a spectrum equivalent to a standard spectrum is acquired. Therefore, a sound signal equivalent to what collected the sound with the standard microphone is acquired.

[0031] in addition, a complement sound -- a user -- although the composite tone acquired as a result of being compounded, when belonging to him also becomes his thing, even if based on data with a standard complement sound, since the part of a complement sound is only a part of a high region, there is very little sense of incongruity.

[0032] In addition, as shown in drawing 2 , it is possible to have individual data memory 3d, to record [voice / of a user] beforehand (sampling), and to store the data of the description part to generate the lack spectrum according to a user's voice.

[0033] Moreover, what the quality of a user's voice is judged and chooses a near thing automatically out of the standard spectrum with which lack spectrum generating was equipped is possible.

<the effectiveness acquired by the configuration (1)> -- according to the above voice pickup systems, the clear sound signal near the usual voice can be generated by carrying out speech recognition and complementing a lack spectrum, after detecting a bone conduction sound and a respiratory tract sound, without spoiling the description of his voice. Moreover, it is not influenced of the surrounding noise.

[0034] Moreover, about a meaningless sound, although detected by pickup like *****, since semantics is not made by processing of speech recognition, a complement sound is not generated. Therefore, since only the low-pass component detected by pickup is outputted, there are few bad influences.

[0035] <The voice pickup structure of a system (2)> The voice pickup structure of a system which is the second example of this invention is first explained using drawing 3 .

[0036] Pickup 1 is a detection means to detect a bone conduction sound or a respiratory tract sound, and a respiratory tract sound microphone corresponds as what detects a bone conduction microphone and a respiratory tract sound as what detects a bone conduction sound.

[0037] The speech recognition circuit 2 is a circuit which recognizes well-known voice or well-known syllable, and whether it learns each description part of a user or it does not perform such study, it is not cared about.

[0038] The artificial sound generating circuit 5 generates the artificial sound corresponding to the sound detected by pickup 1 using the data of the sound (voice, syllable) recognized in said speech recognition circuit 2.

[0039] Actuation of the voice pickup system which is one example of <actuation

[of a voice pickup system] (2)> this invention is constituted by each step of **, **, and ** as roughly divided and shown below. Order is explained for this step later on.

[0040] ** Detection of a sound (a bone conduction sound, respiratory tract sound) : detect a bone conduction sound or a respiratory tract sound using pickup 1.

** Recognition of a sound (a bone conduction sound, respiratory tract sound) : recognize the detection sound of pickup 1 in the speech recognition circuit 2. In this case, you may be any although recognition by the single sound or recognition by syllable changes with recognition approaches of the speech recognition circuit 2. Moreover, it does not matter whether it learns and recognizes a user's voice, or it recognizes without performing study which specified the user.

[0041] ** Spectrum generating of a recognition sound : generate the standard spectrum which will be obtained when the artificial sound generating circuit 5 detects with the usual microphone using the data of a recognition sound. For this reason, the spectrum which the artificial sound generating circuit 5 shall have the spectrum according to a recognition sound (a single sound, syllable), and corresponds according to a recognition sound is called.

[0042] As a spectrum in this case, you may have a spectrum for every user and may have a standard spectrum. Moreover, when it has a standard spectrum, the configuration which changes and uses them like an adult male, an adult woman, and a child as has some spectrums may be used. Therefore, a sound signal equivalent to what collected the sound with the standard microphone is acquired.

[0043] In addition, as shown in drawing 4 , it is possible to have the individual data memory 6, to record [voice / of a user] beforehand (sampling), and to store the data of the description part to generate the lack spectrum according to a user's voice.

[0044] Moreover, what the quality of a user's voice is judged and chooses a near thing automatically out of the standard spectrum with which lack spectrum generating was equipped is possible.

<the effectiveness acquired by the configuration (2)> -- according to the above voice pickup systems, the clear sound signal near the usual voice can be generated now by generating the artificial voice which detects a bone conduction sound and a respiratory tract sound, and corresponds with the data by which speech recognition was carried out. Moreover, it is not influenced of the surrounding noise.

[0045] Moreover, about a meaningless sound, although detected by pickup like *****, since semantics is not made by processing of speech recognition, an artificial sound is not generated. Therefore, since it is not outputted, there are few bad influences.

[0046] <other desirable examples> -- although various kinds of application is possible for the above voice pickup systems, it is possible to use it under a noise environment, including in the various systems which transmit voice. For example, clear transmission is attained by including in devices, such as a cellular phone. And it also has the advantage of not being heard against a surrounding sound.

[0047]

[Effect of the Invention] According to the above voice pickup systems, the clear sound signal near the usual voice can be generated by carrying out speech recognition and complementing a lack spectrum, after detecting a bone conduction sound and a respiratory tract sound, without spoiling the description of his voice. Moreover, it is not influenced of the surrounding noise. Therefore, the voice pickup system which can transmit voice in the condition that articulation is high can be realized, without gathering the surrounding noise.

[0048] Moreover, the clear sound signal near the usual voice can be generated now by generating the artificial voice which detects a bone conduction sound and a respiratory tract sound, and corresponds with the data by which speech recognition was carried out. Moreover, it is not influenced of the surrounding noise. Therefore, the voice pickup system which can transmit voice in the condition that articulation is high can be realized, without gathering the surrounding noise.

[Translation done.]

* NOTICES *

JPO and NCIP are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. **** shows the word which can not be translated.

3. In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is the block diagram showing the voice pickup structure of a system of one example of this invention.

[Drawing 2] It is the block diagram showing the configuration of the modification of the voice pickup system of one example of this invention.

[Drawing 3] It is the block diagram showing the voice pickup structure of a system of the second example of this invention.

[Drawing 4] It is the block diagram showing the configuration of the modification of the voice pickup system of the second example of this invention.

[Drawing 5] It is the property Fig. showing an audio spectrum typically.

[Drawing 6] It is the property Fig. showing the observation result of an audio spectrum.

[Description of Notations]

1 Pickup

2 Speech Recognition Circuit

3 Complement Sound Generation Circuit

3a Spectrum generating circuit

3b Comparator circuit

3c Lack spectrum generating circuit

4 Synthetic Circuit

[Translation done.]

* NOTICES *

JPO and NCIP are not responsible for any damages caused by the use of this translation.

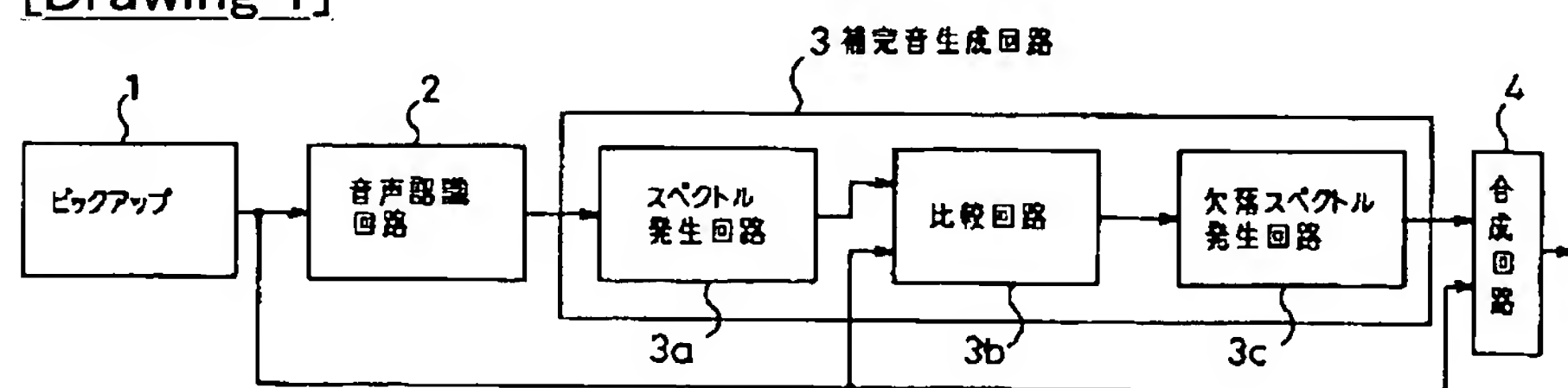
1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.*** shows the word which can not be translated.

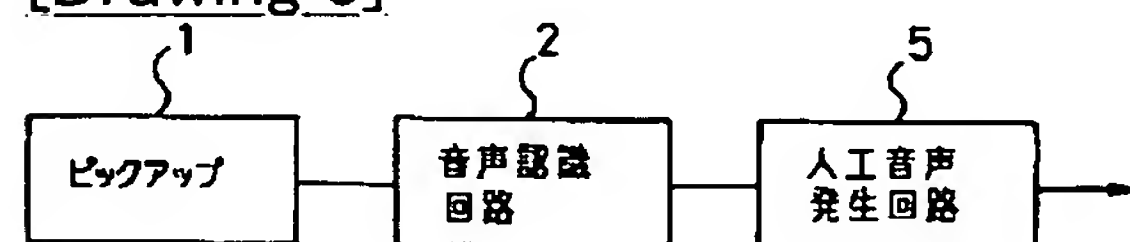
3.In the drawings, any words are not translated.

DRAWINGS

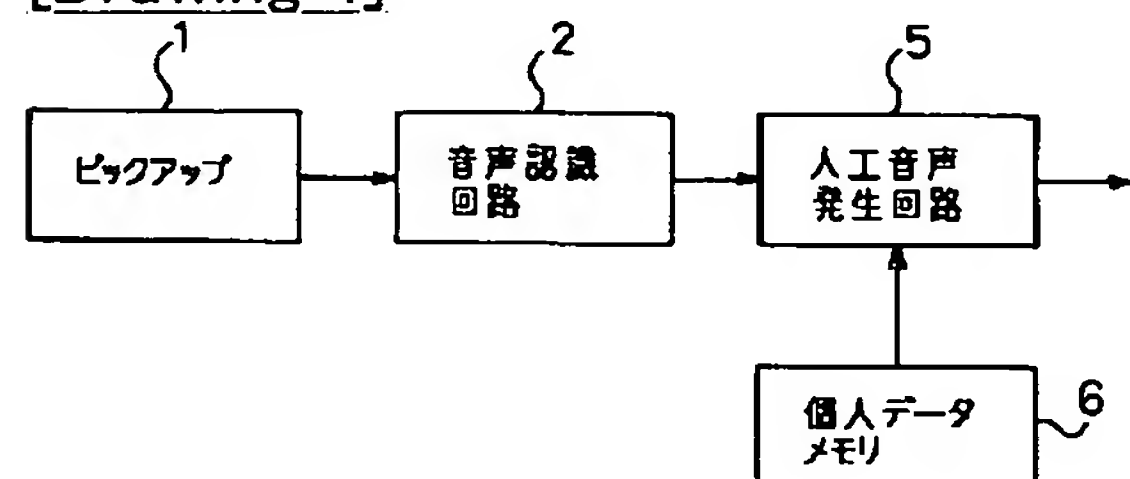
[Drawing 1]



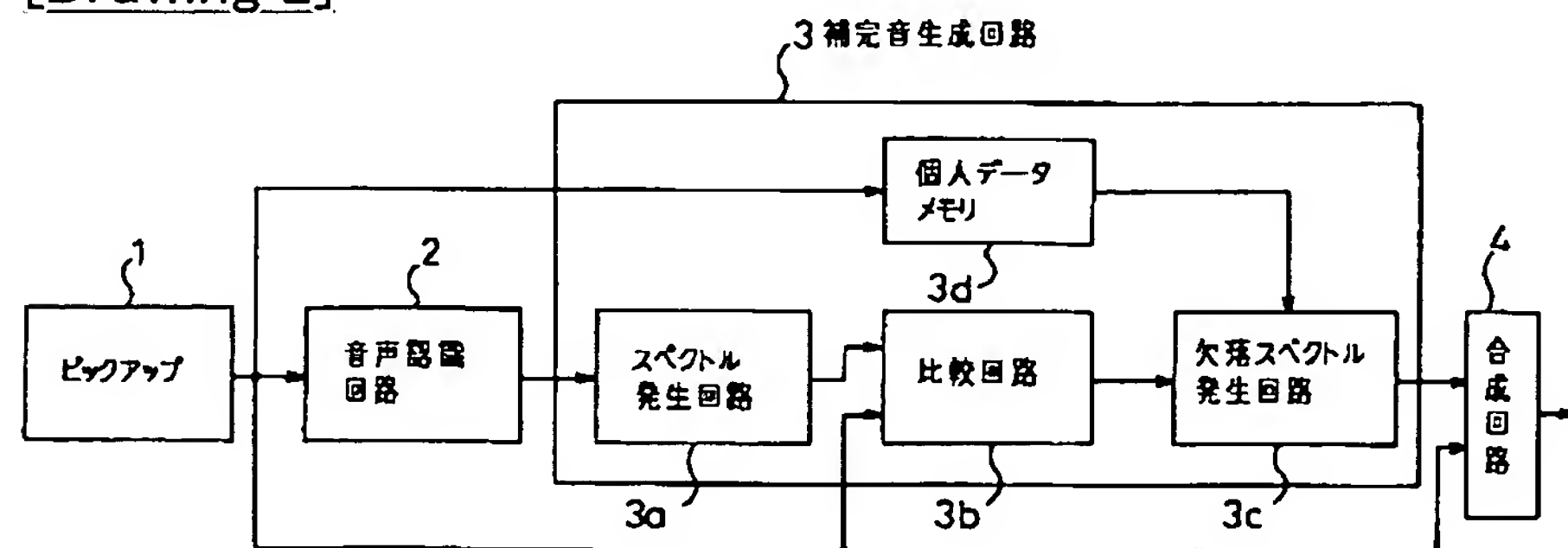
[Drawing 3]



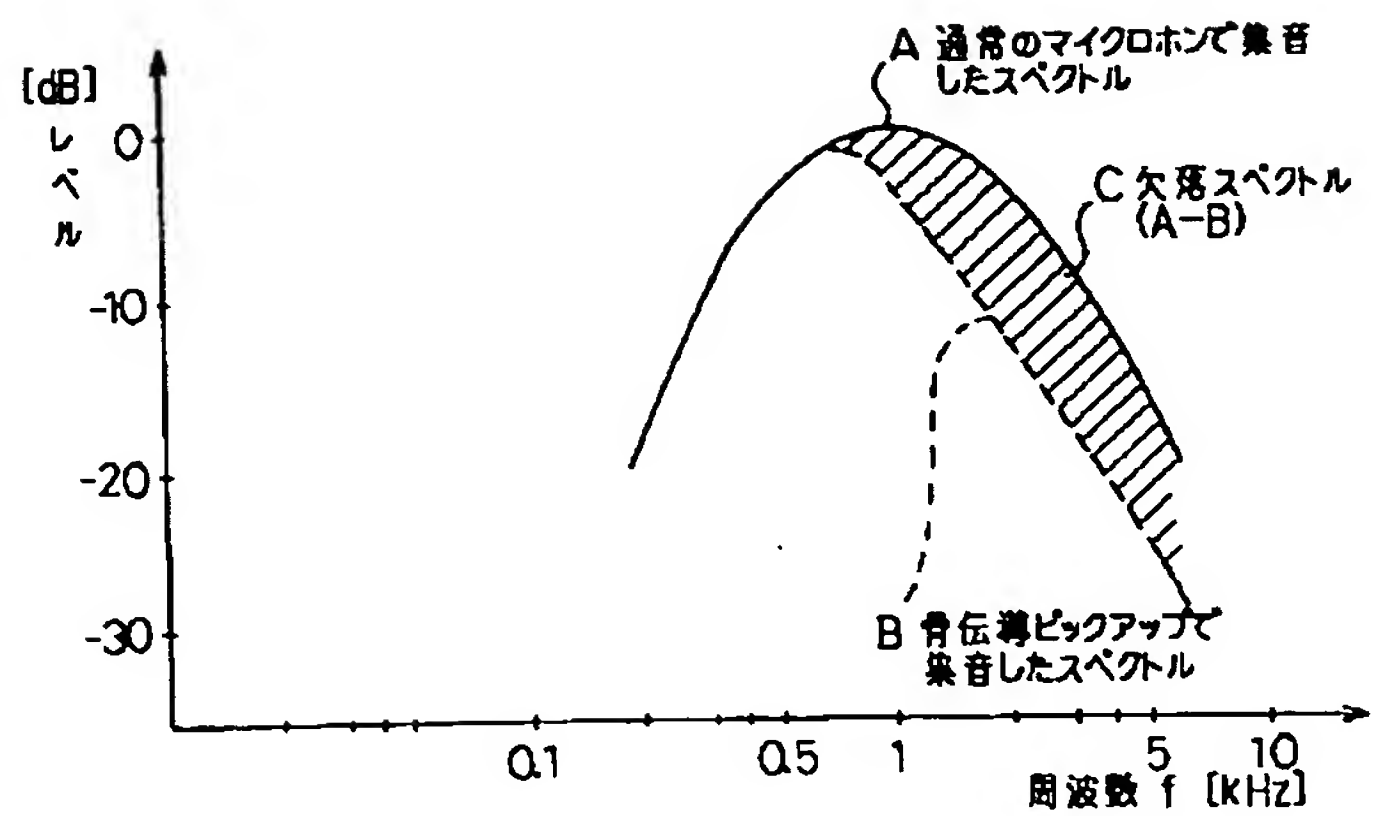
[Drawing 4]



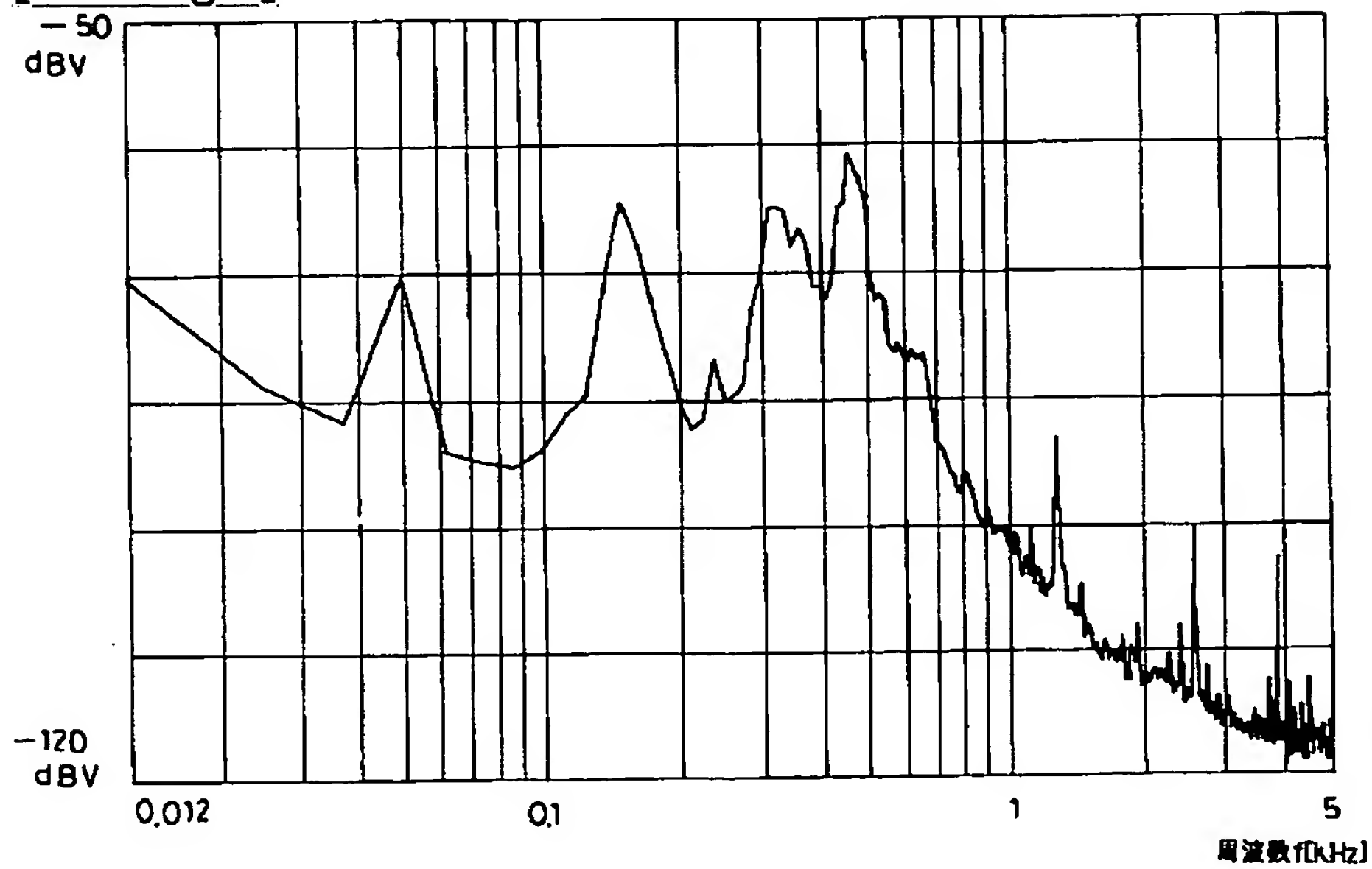
[Drawing 2]



[Drawing 5]



[Drawing 6]



[Translation done.]

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平8-275279

(43)公開日 平成8年(1996)10月18日

(51)Int.Cl. ⁹	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 R 1/00	3 2 7		H 0 4 R 1/00	3 2 7 Z
G 1 0 L 3/00			G 1 0 L 3/00	Q
	5 5 1			5 5 1 C
H 0 4 R 3/00	3 2 0		H 0 4 R 3/00	3 2 0

審査請求 未請求 請求項の数2 O L (全 6 頁)

(21)出願番号 特願平7-75875

(22)出願日 平成7年(1995)3月31日

(71)出願人 000112565

フォスター電機株式会社

東京都昭島市宮沢町512番地

(72)発明者 山口 卓郎

東京都昭島市宮沢町512番地 フォスター
電機株式会社内

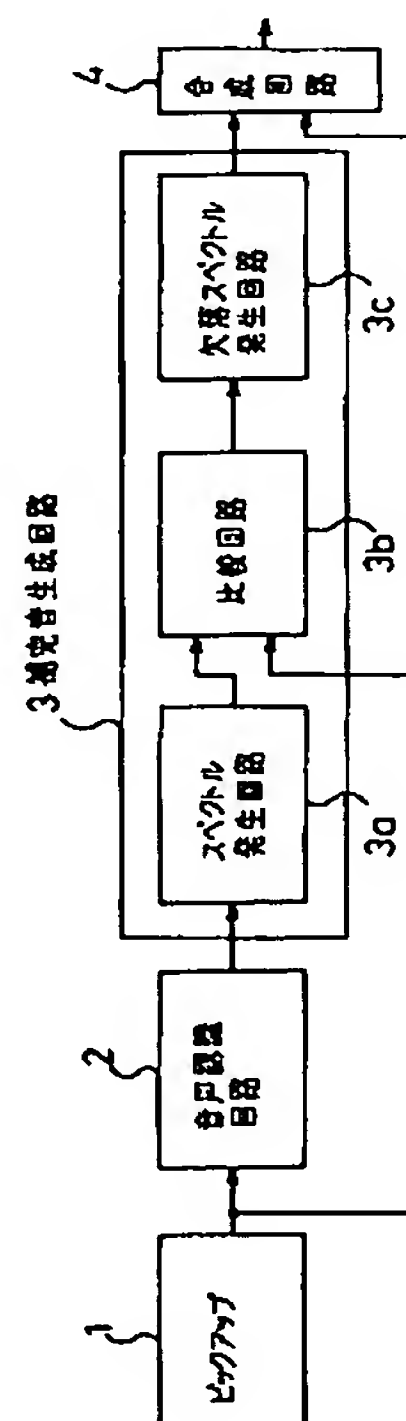
(74)代理人 弁理士 井島 藤治 (外1名)

(54)【発明の名称】 音声ピックアップシステム

(57)【要約】

【目的】 周囲の騒音を拾うことなく、かつ、明瞭度の高い状態で音声を伝達することが可能な音声ピックアップシステムを実現する。

【構成】 骨伝導音若しくは気道音をピックアップ1で検出する音声ピックアップシステムにおいて、前記ピックアップで検出した音を認識可能な音声認識手段2と、前記音声認識手段2で認識された音のデータを用いて、この認識された音のデータを通常のマイクロホンで検出した場合の周波数成分のスペクトルを生成するスペクトル発生手段3aと、このスペクトルと前記ピックアップで検出した音のスペクトルとを比較して欠落している周波数成分を求める比較手段3bと、前記比較手段3bでの比較の結果から前記ピックアップで検出した音に欠落している周波数成分の音を補完音として生成する欠落スペクトル発生手段3cと、前記ピックアップ1で検出した音と前記欠落スペクトル発生手段3cで生成した補完音とを合成して出力する合成手段4とを有する音声ピックアップシステム。



【特許請求の範囲】

【請求項1】 骨伝導音若しくは気道音をピックアップで検出する音声ピックアップシステムにおいて、前記ピックアップで検出した音を認識可能な音声認識手段と、

前記音声認識手段で認識された音のデータを用いて、この認識された音のデータを通常のマイクロホンで検出した場合の周波数成分のスペクトルを生成するスペクトル発生手段と、

このスペクトル発生手段が生成したスペクトルと前記ピックアップで検出した音のスペクトルとを比較して欠落している周波数成分を求める比較手段と、

前記比較手段での比較の結果から前記ピックアップで検出した音に欠落している周波数成分の音を補完音として生成する欠落スペクトル発生手段と、

前記ピックアップで検出した音と前記欠落スペクトル発生手段で生成した補完音とを合成して出力する合成手段と、

を有することを特徴とする音声ピックアップシステム。

【請求項2】 骨伝導音若しくは気道音をピックアップで検出する音声ピックアップシステムにおいて、前記ピックアップで検出した音を認識可能な音声認識手段と、

前記音声認識手段で認識された音のデータに対応する音を人工音として生成する人工音生成手段と、

を有することを特徴とする音声ピックアップシステム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は音声ピックアップシステムに関し、更に詳しくは、骨伝導音や気道音をピックアップで検出する音声ピックアップシステムに関する。

【0002】

【従来の技術】騒音が存在する環境で使用するマイクロホンとして、骨伝導音を検出する骨伝導音ピックアップが知られている。

【0003】この骨伝導音ピックアップはユーザの顔面や頭部に密着させておき、ユーザの発声に伴う顔面や頭部の振動を検出するものである。また、似たようなマイクロホンとして、イヤホンのような形状のピックアップを外耳に挿入して、外耳道の気道音を検出するものも存在している。

【0004】以上のような骨伝導音や気道音を検出するピックアップは周囲の騒音を比較的拾わずに、目的とするユーザの音声を検出し易いという利点を有する。

【0005】

【発明が解決しようとする課題】しかし、通常の音声のスペクトルが300Hz～3kHzであるとした場合に、図5特性Bに示すように、1kHz以上の周波数領域で検出レベルが低下する問題を有している。

【0006】すなわち、骨伝導音や気道音といった間接

的に音声を検出する形式のピックアップの検出音は、通常のマイクロホンで検出した音声信号（図5特性A）と比較して、音声の低域成分が強調されて高域成分が徐々に低下する感じになり（図5特性B）、明瞭度が低下する問題を有している。

【0007】図6は男性の声を実際に骨伝導音ピックアップで検出した場合の周波数特性を示す特性図である。この特性図からも高域成分の低下の様子が読み取れる。実際には、骨伝導音ピックアップを顔面若しくは頭部に押さえつける際の圧力や、男性／女性の別などによって高域成分の低下の度合は若干異なるが、高域成分が低下することには変わりがない。

【0008】本発明は上記の問題点に鑑みてなされたもので、その目的は、周囲の騒音を拾うことなく、かつ、明瞭度の高い状態で音声を伝達することが可能な音声ピックアップシステムを提供することにある。

【0009】

【課題を解決するための手段】本件出願の発明者は、従来の音声ピックアップシステムにおいて予想される明瞭度等の不具合を改良すべく鋭意研究を行った結果、従来は周波数特性の点で明瞭度に問題を有していた骨伝導音や気道音を検出するピックアップにおいても明瞭度の高い音声を伝達できる構成を見出し、本発明を完成させたものである。

【0010】従って、課題を解決する手段である本発明は以下に説明するように構成されたものである。

（1）すなわち、上記の課題を解決する第1の手段は、骨伝導音若しくは気道音をピックアップで検出する音声ピックアップシステムにおいて、前記ピックアップで検出した音を認識可能な音声認識手段と、前記音声認識手段で認識された音のデータを用いて、この認識された音のデータを通常のマイクロホンで検出した場合の周波数成分のスペクトルを生成するスペクトル発生手段と、このスペクトル発生手段が生成したスペクトルと前記ピックアップで検出した音のスペクトルとを比較して欠落している周波数成分を求める比較手段と、前記比較手段での比較の結果から前記ピックアップで検出した音に欠落している周波数成分の音を補完音として生成する欠落スペクトル発生手段と、前記ピックアップで検出した音と前記欠落スペクトル発生手段で生成した補完音とを合成して出力する合成手段と、を有することを特徴とする音声ピックアップシステムである。

【0011】尚、このような音声ピックアップシステムにおいて、欠落スペクトル発生手段において生成する補完音としては、予めシステムに音声合成用の学習をさせておいて、使用者の音声に似せた音声を発生するために必要な欠落スペクトルとすることも可能である。

【0012】（2）また、上記の課題を解決する第2の手段は、骨伝導音若しくは気道音をピックアップで検出する音声ピックアップシステムにおいて、前記ピックア

ップで検出した音を認識可能な音声認識手段と、前記音声認識手段で認識された音のデータに対応する音を人工音として生成する人工音生成手段と、を有することを特徴とする音声ピックアップシステムである。

【0013】尚、このような音声ピックアップシステムにおいて、人工音生成手段において生成する人工音としては、予めシステムに音声合成用の学習をさせておいて使用者の音声に似せた音声を発生することも、別の一般的な音声を発生することも可能である。

【0014】

【作用】課題を解決する第1の手段である音声ピックアップシステムにおいて、骨伝導音若しくは気道音をピックアップで検出し、ピックアップで検出した音を音声認識手段で認識し、音声認識手段で認識された音のデータを用いて、この認識された音のデータを通常のマイクロホンで検出した場合の周波数成分のスペクトルを生成し、このように生成したスペクトルとピックアップで実際に検出した音のスペクトルとを比較手段で比較し欠落している周波数成分を求めて、比較手段での比較の結果からピックアップで検出した音に欠落している周波数成分の音を欠落スペクトル発生手段で補完音として生成し、ピックアップで検出した音と欠落スペクトル発生手段で生成した補完音とを合成手段において合成して出力する。

【0015】以上のような音声ピックアップシステムによれば、骨伝導音や気道音を検出した後に音声認識して欠落スペクトルを補完することで、本人の音声の特徴を損なうことなく通常の音声に近い明瞭な音声信号を生成することができるようになる。また、周囲の騒音の影響を受けることもない。

【0016】課題を解決する第2の手段である音声ピックアップシステムにおいて、骨伝導音若しくは気道音をピックアップで検出し、ピックアップで検出した音を音声認識手段で認識し、音声認識手段で認識された音に対応する音を人工音生成手段で人工音として生成する。

【0017】以上のような音声ピックアップシステムによれば、骨伝導音や気道音を検出し、音声認識されたデータにより対応する人工音声を発生することで、通常の音声に近い明瞭な音声信号を生成することができるようになる。また、周囲の騒音の影響を受けることもない。

【0018】

【実施例】図面を用いて本発明の一実施例について詳細に説明する。

<音声ピックアップシステムの構成(1)>まず、本発明の一実施例である音声ピックアップシステムの構成について図1を用いて説明を行なう。

【0019】ピックアップ1は骨伝導音若しくは気道音などを検出する検出手段であり、骨伝導音を検出するものとしては骨伝導マイクロホン(骨伝導音ピックアップ)、気道音を検出するものとしては気道音マイクロホ

ンが該当する。

【0020】音声認識回路2は周知の音声若しくは音節を認識する回路であり、ユーザの個々の特徴部分を学習するものであっても、また、このような学習を行わないものであっても構わない。

【0021】補完音生成回路3は前記音声認識回路2で認識された音(音声、音節)のデータを用いて、ピックアップ1で検出された音に欠落している周波数成分の音を補完音として生成する一種の人工音発生回路である。

10 また、この補完音生成回路3は、スペクトル発生回路3aと、比較回路3bと、欠落スペクトル発生回路3cとから構成されている。

【0022】スペクトル発生回路3aは前記音声認識回路2で認識された音のデータを用いて、認識された音が通常のマイクロホンで検出された場合のスペクトルを発生する。比較回路3bは前記スペクトル発生回路3aが発生したスペクトルと、前記ピックアップで検出した音のスペクトルとを比較する。欠落スペクトル発生回路3cは、比較回路3bの比較結果に応じて、スペクトルの
20 差分に応じた部分のスペクトル(欠落スペクトル)を補完音として発生する。

【0023】合成回路4はピックアップ1で検出された音(骨伝導音、気道音)と補完音生成回路3で生成された補完音とを合成して出力する出力手段である。

<音声ピックアップシステムの動作(1)>本発明の一実施例である音声ピックアップシステムの動作は、大きく分けて以下に示したような①、②、③、④、⑤、⑥の各ステップにより構成されている。このステップを順を追って説明する。

30 【0024】①音(骨伝導音、気道音)の検出:ピックアップ1を用いて骨伝導音若しくは気道音を検出する。
②音(骨伝導音、気道音)の認識:ピックアップ1の検出音を音声認識回路2で認識する。この場合、音声認識回路2の認識方法により、単音での認識か音節での認識かが異なるが、いずれであっても構わない。また、ユーザの音声を学習して認識するものであっても、また、ユーザを特定した学習を行わないで認識するものであっても構わない。

40 【0025】③認識音のスペクトル発生:認識音のデータを用いて、スペクトル発生回路3aが通常のマイクロホンで検出した場合に得られるであろうスペクトル(以下、これを標準音のスペクトルと言う)を発生する。このために、スペクトル発生回路3aは認識音(単音、音節)に従ったスペクトルを有しているものとし、認識音に従って対応するスペクトルが呼び出されるようになっている。

50 【0026】この場合のスペクトルとしては、ユーザ毎のスペクトルを有しても良いし、標準的なスペクトルを有しても良い。また、標準的なスペクトルを有するとした場合には、成人男性、成人女性、子供等のように幾つ

かのスペクトルを有するようにしても構わない。

【0027】④標準音と検出音とのスペクトル比較：比較回路3bにおいて、標準音のスペクトルと検出音とのスペクトルとを比較する。

【0028】例えば、図5を用いて説明すると、認識された検出音毎に、標準音のスペクトルAと検出音のスペクトルBとを比較して、検出音の欠落スペクトルC(=A-B)を算出する。

【0029】⑤補完音(欠落スペクトル)発生：比較回路3bで得られた欠落スペクトルのデータに応じて欠落スペクトル発生回路3cが欠落スペクトルの信号を発生する。この場合も、スペクトル発生回路3aと同じ様に、発生する欠落スペクトルとして、ユーザ毎のスペクトルを有しても良いし、標準的なスペクトルを有しても良い。また、標準的なスペクトルを有するとした場合には、成人男性、成人女性、子供等のように幾つかのスペクトルを有するようにしても構わない。

【0030】⑥検出音と補完音との合成：合成回路4において、欠落スペクトルと検出音のスペクトルとを合成する。この合成処理により、検出音の欠落スペクトルが補完音として加算され、標準スペクトルと同等なスペクトルの合成音が得られる。従って、標準的なマイクロホンで集音したものと同等な音声信号が得られる。

【0031】尚、補完音がユーザ本人のものであれば合成された結果得られる合成音も本人のものとなるが、補完音が標準的なデータに基づくものであったとしても、補完音の部分は高域の部分のみであるので違和感は極めて少ない。

【0032】尚、ユーザの声に応じた欠落スペクトルを発生したい場合には、図2に示すように、個人データメモリ3dを備えておいて、ユーザの音声を予め収録(サンプリング)しておいて特徴部分のデータを格納しておくことが可能である。

【0033】また、ユーザの声の質を判定して、欠落スペクトル発生用に複数備えた標準的なスペクトルの中から近いものを自動的に選択するようなことも可能である。

<構成(1)により得られる効果>以上のような音声ピックアップシステムによれば、骨伝導音や気道音を検出した後に音声認識して欠落スペクトルを補完することで、本人の音声の特徴を損なうことなく通常の音声に近い明瞭な音声信号を生成することができるようになる。また、周囲の騒音の影響を受けることもない。

【0034】また、歯噛み音などのようにピックアップで検出されるものの無意味な音については、音声認識の処理で意味をなさないので補完音が生成されない。従って、ピックアップで検出された低域成分のみが出力されるため、悪影響は少ない。

【0035】<音声ピックアップシステムの構成(2)>まず、本発明の第二の実施例である音声ピックアップ

システムの構成について図3を用いて説明を行なう。

【0036】ピックアップ1は骨伝導音若しくは気道音などを検出する検出手段であり、骨伝導音を検出するものとしては骨伝導マイクロホン、気道音を検出するものとしては気道音マイクロホンが該当する。

【0037】音声認識回路2は周知の音声若しくは音節を認識する回路であり、ユーザの個々の特徴部分を学習するものであっても、また、このような学習を行わないものであっても構わない。

10 【0038】人工音発生回路5は前記音声認識回路2で認識された音(音声、音節)のデータを用いて、ピックアップ1で検出された音に対応した人工音を生成するものである。

【0039】<音声ピックアップシステムの動作(2)>本発明の一実施例である音声ピックアップシステムの動作は、大きく分けて以下に示したような①、②、③の各ステップにより構成されている。このステップを順を追って説明する。

20 【0040】①音(骨伝導音、気道音)の検出：ピックアップ1を用いて骨伝導音若しくは気道音を検出する。
②音(骨伝導音、気道音)の認識：ピックアップ1の検出音を音声認識回路2で認識する。この場合、音声認識回路2の認識方法により、単音での認識か音節での認識かが異なるが、いずれであっても構わない。また、ユーザの音声を学習して認識するものであっても、また、ユーザを特定した学習を行わないで認識するものであっても構わない。

30 【0041】③認識音のスペクトル発生：認識音のデータを用いて、人工音発生回路5が通常のマイクロホンで検出した場合に得られるであろう標準スペクトルを発生する。このために、人工音発生回路5は認識音(単音、音節)に従ったスペクトルを有しているものとし、認識音に従って対応するスペクトルが呼び出されるようになっている。

40 【0042】この場合のスペクトルとしては、ユーザ毎のスペクトルを有しても良いし、標準的なスペクトルを有しても良い。また、標準的なスペクトルを有するとした場合には、成人男性、成人女性、子供等のように幾つかのスペクトルを有するようにして切り替えて使用する構成でも構わない。従って、標準的なマイクロホンで集音したものと同等な音声信号が得られる。

【0043】尚、ユーザの声に応じた欠落スペクトルを発生したい場合には、図4に示すように、個人データメモリ6を備えておいて、ユーザの音声を予め収録(サンプリング)しておいて特徴部分のデータを格納しておくことが可能である。

50 【0044】また、ユーザの声の質を判定して、欠落スペクトル発生用に複数備えた標準的なスペクトルの中から近いものを自動的に選択するようなことも可能である。

＜構成（２）により得られる効果＞以上のような音声ピックアップシステムによれば、骨伝導音や気道音を検出し、音声認識されたデータにより対応する人工音声を発生することで、通常の音声に近い明瞭な音声信号を生成することができるようになる。また、周囲の騒音の影響を受けることもない。

【００４５】また、歯噛み音などのようにピックアップで検出されるものの無意味な音については、音声認識の処理で意味をなさないため人工音が生成されない。従って、出力されないため悪影響は少ない。

【００４６】＜その他の好ましい例＞以上のような音声ピックアップシステムは各種の応用が可能であるが、騒音環境下で音声を伝達する各種システムに組み込んで使用することが可能である。例えば、携帯電話等の機器に組み込むことで明瞭な送話が可能になる。そして、周囲の音を相手に聞かれることが無いという利点も有している。

【００４７】

【発明の効果】以上のような音声ピックアップシステムによれば、骨伝導音や気道音を検出した後に音声認識して欠落スペクトルを補完することで、本人の音声の特徴を損なうことなく通常の音声に近い明瞭な音声信号を生成することができるようになる。また、周囲の騒音の影響を受けることもない。従って、周囲の騒音を拾うことなく、かつ、明瞭度の高い状態で音声を伝達することが可能な音声ピックアップシステムを実現できるようになる。

【００４８】また、骨伝導音や気道音を検出し、音声認

識されたデータにより対応する人工音声を発生することで、通常の音声に近い明瞭な音声信号を生成することができるようになる。また、周囲の騒音の影響を受けることもない。従って、周囲の騒音を拾うことなく、かつ、明瞭度の高い状態で音声を伝達することが可能な音声ピックアップシステムを実現できるようになる。

【図面の簡単な説明】

【図１】本発明の一実施例の音声ピックアップシステムの構成を示す構成図である。

10 【図２】本発明の一実施例の音声ピックアップシステムの変形例の構成を示す構成図である。

【図３】本発明の第二の実施例の音声ピックアップシステムの構成を示す構成図である。

【図４】本発明の第二の実施例の音声ピックアップシステムの変形例の構成を示す構成図である。

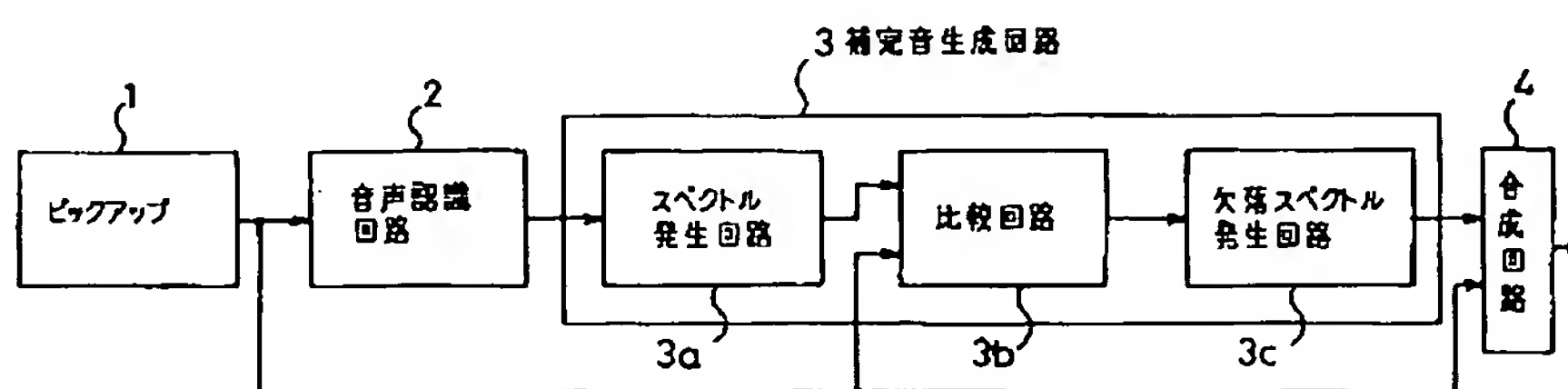
【図５】音声のスペクトルを模式的に示す特性図である。

【図６】音声のスペクトルの実測結果を示す特性図である。

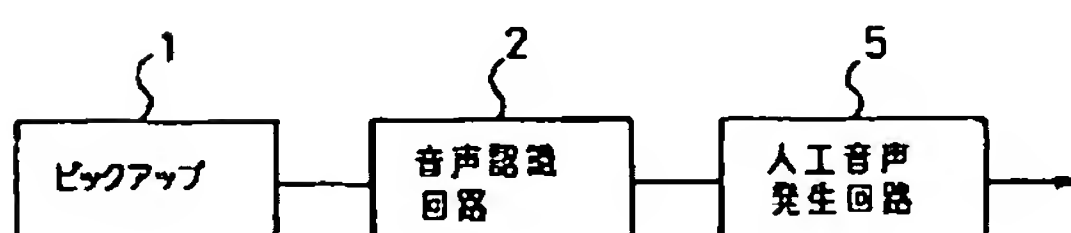
20 【符号の説明】

- １ ピックアップ
- ２ 音声認識回路
- ３ 補完音生成回路
- ３a スペクトル発生回路
- ３b 比較回路
- ３c 欠落スペクトル発生回路
- ４ 合成回路

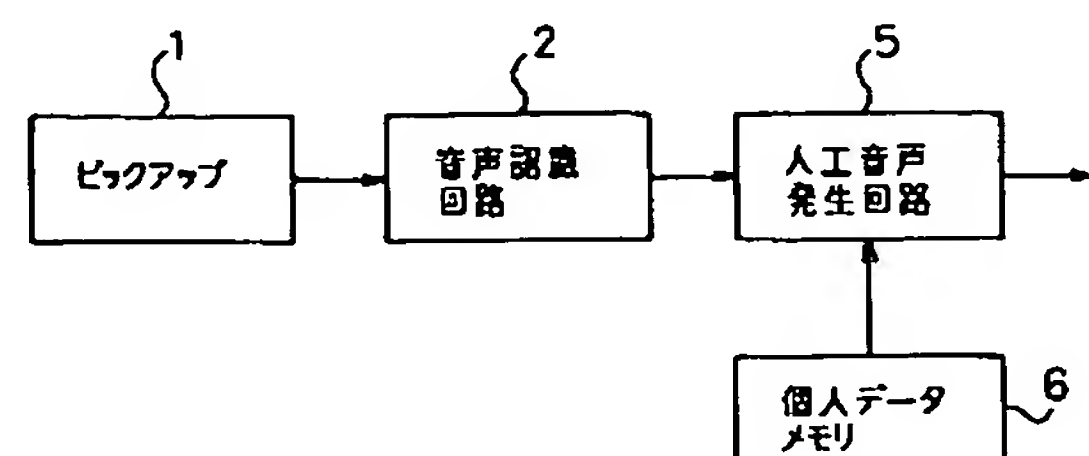
【図１】



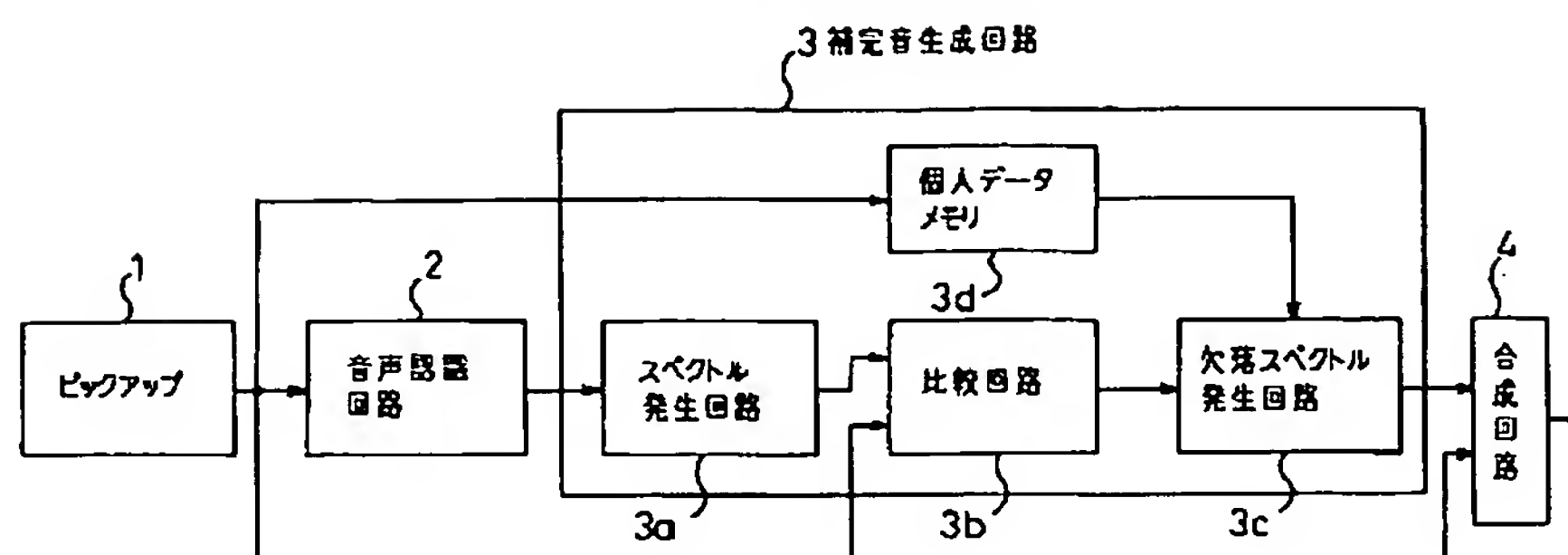
【図３】



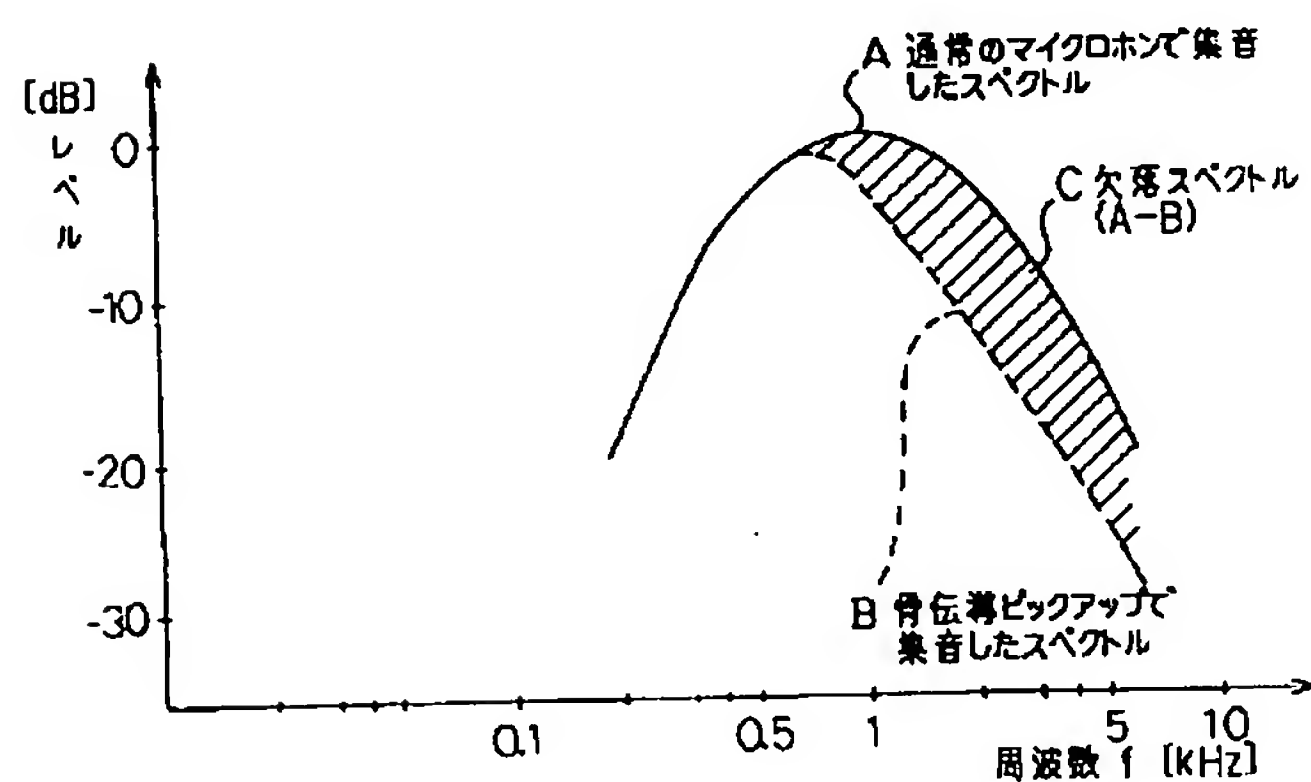
【図４】



【図 2】



【図 5】



【図 6】

